Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 1 February 2008 (MN LATEX style file v2.2)

Recovering galaxy star formation and metallicity histories from spectra using VESPA

R. Tojeiro⋆[1], A. F. Heavens[1], R. Jimenez[2] and B. Panter[1]
[1] *Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK*
[2] *Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA-19104, USA*

1 February 2008

ABSTRACT
We introduce VErsatile SPectral Analysis (VESPA): a new method which aims to recover robust star formation and metallicity histories from galactic spectra. VESPA uses the full spectral range to construct a galaxy history from synthetic models. We investigate the use of an adaptative parametrization grid to recover reliable star formation histories on a galaxy-by-galaxy basis. Our goal is robustness as opposed to high resolution histories, and the method is designed to return high time resolution only where the data demand it. In this paper we detail the method and we present our findings when we apply VESPA to synthetic and real Sloan Digital Sky Survey (SDSS) spectroscopic data. We show that the number of parameters that can be recovered from a spectrum depends strongly on the signal-to-noise, wavelength coverage and presence or absence of a young population. For a typical SDSS sample of galaxies, we can normally recover between 2 to 5 stellar populations. We find very good agreement between VESPA and our previous analysis of the SDSS sample with MOPED.

Key words: methods: data analysis - methods: statistical - galaxies: stellar content - galaxies: evolution - galaxies: formation

1 INTRODUCTION

The spectrum of a galaxy holds vasts amounts of information about that galaxy's history and evolution. Finding a way to tap directly into this source of knowledge would not only provide us with crucial information about that galaxy's evolutionary path, but would also allow us to integrate this knowledge over a large number of galaxies and therefore derive cosmological information.

Galaxy formation and evolution are still far from being well understood. Galaxies are extremely complex objects, formed via complicated non-linear processes, and any approach (be it observational, semi-analytical or computational) inevitably relies on simplifications. If we try to analyse a galaxy's luminous output in terms of a history parametrized by some chosen physical quantities, such a simplification is also in order. The reason is two-fold: firstly we are limited by our knowledge and ability to model all the physical processes which happen in a galaxy and produce the observed spectrum we are analysing; secondly, the observed spectrum is inevitably perturbed by noise, which intrinsically limits the amount of information we can recover.

Measuring and understanding the star formation history of the Universe is therefore essential to our understanding of galaxy evolution - when, where and in what conditions did stars form throughout cosmic history? The traditional and simplest way to probe this is to measure the observed instantaneous star formation rate in galaxies at different redshifts. This can be achieved by looking at light emitted by young stars in the ultra-violet (UV) band or its secondary effects. (e.g. Madau et al. 1996; Kennicutt 1998; Hopkins et al. 2000; Bundy et al. 2006; Erb et al. 2006; Abraham et al. 2007; Noeske et al. 2007; Verma et al. 2007). A complementary method is to look at present day galaxies and extract their star formation history, which spans the lifetime of the galaxy. Different teams have analysed a large number of galaxies in this way, whether by using the full available spectrum (Glazebrook et al. 2003; Panter et al. 2003; Cid Fernandes et al. 2004; Heavens et al. 2004; Mathis et al. 2006; Ocvirk et al. 2006; Panter et al. 2006; Cid Fernandes et al. 2007), or by concentrating on particular spectral features or indices (e.g. Kauffmann et al. 2003; Tremonti et al. 2004; Gallazzi et al. 2005; Barber et al. 2006), which are known to be correlated with age or metallicities (e.g. Worthey 1994; Thomas et al.

⋆ E-mail: rmft@roe.ac.uk

2003).

To do this, we rely on synthetic stellar population models to describe a galaxy in terms of its stellar components, but by modelling a galaxy in this way we are intrinsically limited by the quality of the models. There are also potential concerns with flux calibration errors. However, using the full spectrum to recover the fossil record of a galaxy - or of an ensemble of galaxies - is an extremely powerful method, as the quality and amount of data relating to local galaxies vastly outshines that which concerns high-redshift galaxies. Splitting a galaxy into simple stellar populations of different ages and metallicities is a natural way of parameterising a galaxy, and it allows realistic fits to real galaxies (e.g. Bruzual & Charlot 2003). Galactic archeology has become increasingly popular in the literature recently, largely due to the increase in sophistication of stellar population synthesis codes and the improvement of the stellar spectra libraries upon which they are based, and also due to the availability of large well-calibrated spectroscopic databases, such as the Sloan Digital Sky Survey (SDSS) (York et al. 2000; Strauss et al. 2002).

In any case, without imposing any constraints on the allowed form of the star formation history, or perhaps an age-metallicity relation, the parameter space can become unsustainably large for a traditional approach. Ideally, one would like to do without such pre-constraints. Recently, different research teams have come up with widely different solutions for this problem. MOPED (Heavens et al. 2000) and STARLIGHT (Cid Fernandes et al. 2004) explore a well-chosen parameter space in order to find the best possible fit to the data. In the case of MOPED, this relies on compression of the full spectrum to a much smaller set of numbers which retains all the information about the parameters it tries to recover; STARLIGHT on the other hand, searches for its best fit using the full spectrum with a Metropolis algorithm. STECMAP (Ocvirk et al. 2006) solves the problem using an algebraic least-squares solution and a well-chosen regularization to keep the inversions stable. All of these and other methods acknowledge the same limitation - noise in the data and in the models introduces degeneracies into the problem which can lead to unphysical results. MOPED, for example, has produced some remarkable results concerning the average star formation history of the Universe by analysing a large sample of galaxies. However, MOPED's authors have cautioned against over-interpreting the results on a galaxy-by-galaxy basis, due to the problem mentioned above. This is directly related to the question of how finely one should parameterise a galaxy, and what the consequences of this might be.

Much of the motivation for VESPA came from the realisation that this problem will vary from galaxy to galaxy, and that the method of choosing a single parametrization to analyse a large number of galaxies can be improved on.

VESPA is based on three main ideas, which we present here and develop further in the main text:

- There is only so much information one can safely recover from any given set of data, and the amount of information which can be recovered from an individual galaxy varies.
- The recovered star formation fractions should be positive.
- Even though the full unconstrained problem is non-linear, it is piecewise linear in well-chosen regions of parameter space.

VESPA's ultimate goal is to derive robust information for each galaxy individually, by adapting the number of parameters it recovers on a galaxy-by-galaxy basis and increasing the resolution in parameter space only where the data warrant it. In a nutshell, this is how VESPA works: we estimate how many parameters we can recover from a given spectrum, given its noise, shape, spectral resolution and wavelength range using an analysis given by Ocvirk et al. (2006). In that paper, Singular Value Decomposition (SVD) is used to find a least squares solution, and this solution is analysed in terms of its singular vectors. VESPA uses this method only as an analysis of the solution, and uses Bounded-Variable Least-Squares (BVLS) (Stark & Parker 1995) to reach a non-negative solution in several regimes where linearity applies.

This paper is organised as follows: in Section 2 we present the method, in Section 3 we apply VESPA to a variety of synthetic spectra, in Section 4 we apply VESPA to a sample of galaxies from the Sloan Digital Sky Survey spectroscopic database and we compare our results to those obtained with MOPED, and finally in Section 5 we present our conclusions.

2 METHOD

In this section we lay down the problem to solve in detail, and explain the different steps VESPA uses to find a solution for each galaxy.

2.1 The problem

We assume a galaxy is composed of a series of simple stellar populations (SSP) of varying ages and metallicities. The unobscured rest frame luminosity per unit wavelength of a galaxy can then be written as

$$F_\lambda = \int_0^t \psi(t) S_\lambda(t, Z) dt \tag{1}$$

where $\psi(t)$ is the star formation rate (solar masses formed per unit of time) and $S_\lambda(t, Z)$ is the luminosity per unit wavelength of a single stellar population of age t and metallicity Z, per unit mass. The dependency of the metallicity on age is unconstrained, turning this into a non-linear problem.

In order to solve this problem, we start by discretizing in wavelength and time, by averaging these two quantities into well chosen bins. For now we present the problem with a generalised parametrization, and discuss our choice in Section 2.3. We will use greek indices to indicate time bins, and roman indices to indicate wavelength bins.

The problem becomes

$$F_j = \sum_\alpha x_\alpha G(Z_\alpha)_{\alpha j} \quad (2)$$

where $F_j = (F_1, ..., F_D)$ is the luminosity of the jth wavelength bin of width $\Delta\lambda$, $G(Z_\alpha)_{\alpha j}$ is the jth luminosity point of a stellar population of age $t_\alpha = (t_1, ..., t_S)$ (spanning an age range of Δt_α) and metallicity Z_α, and $x_\alpha = (x_1, ..., x_S)$ is the total mass of population $G(Z)_{\alpha j}$ in the time bin Δt_α.

Although the full metallicity problem is non-linear, interpolating between tabulated values of Z gives a piecewise linear behaviour:

$$G(Z_\alpha)_{\alpha j} = g_\alpha G(Z_{a,\alpha})_{\alpha j} + (1 - g_\alpha) G(Z_{b,\alpha})_{\alpha j}, \quad (3)$$

and the problem then becomes

$$F_j = \sum_\alpha x_\alpha \left[g_\alpha G(Z_{a,\alpha})_{\alpha j} + (1 - g_\alpha) G(Z_{b,\alpha})_{\alpha j} \right] \quad (4)$$

where $G(Z_{a,\alpha})_{\alpha j}$ and $G(Z_{b,\alpha})_{\alpha j}$ are equivalent to $G(Z_\alpha)_{\alpha j}$ as above, but at fixed metallicities Z_a and Z_b, which bound the true Z. If this interpolation matches the models' resolution in Z, then we are not degrading the models in any way.

Solving the problem then requires finding the correct metallicity range. One should not underestimate the complexity this implies - trying all possible combination of consecutive values of Z_a and Z_b in a grid of 16 age bins would lead to a total number of calculations of the order of 10^9, which is unfeasible even in today's fast personal workstations. We work around this problem using an iterative approach, which we describe in Section 2.3.2.

2.1.1 Dust extinction

An important component when describing the luminous output of a galaxy is dust, as different wavelengths are affected in different ways. The simplest possible approach is to use one-parameter dust model, according to which we apply a single dust screen to the combined luminosity of all the galactic components. Equation (1) becomes

$$F_\lambda = f_{dust}(\tau_\lambda) \int_0^t \psi(t) S_\lambda(t, Z) dt \quad (5)$$

where we are assuming the dust extinction is the same for all stars, and characterised by the optical depth, τ_λ.

However, it is also well known that very young stars are likely to be more affected by dust. In an attempt to include this in our modelling, we follow the two-parameter dust model of Charlot & Fall (2000) in which young stars are embebbed in their birth cloud up to a time t_{BC}, when they break free into the inter-stellar medium (ISM):

$$F_\lambda = \int_0^t f_{dust}(\tau_\lambda, t) \psi(t) S_\lambda(t, Z) dt \quad (6)$$

and

$$f_{dust}(\tau_\lambda, t) = \begin{cases} f_{dust}(\tau_\lambda^{ISM}) f_{dust}(\tau_\lambda^{BC}), t \leq t_{BC} \\ f_{dust}(\tau_\lambda^{ISM}), t > t_{BC} \end{cases} \quad (7)$$

where τ_λ^{ISM} is the optical depth of the ISM and τ_λ^{BC} is the optical depth of the birth cloud. Following Charlot & Fall (2000), we take $t_{BC} = 0.03$ Gyrs.

There is a variety of choices for the form of $f_{dust}(\tau_\lambda)$. To model the dust in the ISM, we use the mixed slab model of Charlot & Fall (2000) for low optical depths ($\tau_V \leq 1$), for which

$$f_{dust}(\tau_\lambda) = \frac{1}{2\tau_\lambda} [1 + (\tau_\lambda - 1) \exp(-\tau_\lambda) - \tau_\lambda^2 E_1(\tau_\lambda)] \quad (8)$$

where E_1 is the exponential integral and τ_λ is the optical depth of the slab. This model is known to be less accurate for high dust values, and for optical depths greater than one we take a uniform screening model with

$$f_{dust}(\tau_\lambda) = \exp(-\tau_\lambda). \quad (9)$$

We only use the uniform screening model to model the dust in the birth cloud and we use $\tau_\lambda = \tau_V (\lambda / 5500\mathring{A})^{-0.7}$ as our extinction curve for both environments.

As described, dust is a non-linear problem. In practice, we solve the linear problem described by equation (4) with a number of dust extinctions applied to the matrices $G(Z)_{ij}$ and choose the values of τ_V^{ISM} and τ_V^{BC} which result in the best fit to the data.

We initially use a binary chop search for $\tau_V^{ISM} \in [0, 4]$ and keep τ_V^{BC} fixed and equal to zero, which results in trying out typically around nine values of τ_V^{ISM}. If this initial solution reveals star formation at a time less than t_{BC} we repeat our search on a two-dimensional grid, and fit for τ_V^{ISM} and τ_V^{BC} simultaneously. There is no penalty except in CPU time to apply the two-parameter search, but we find that this procedure is robust (see section 3.4).

2.2 The solution

In this section we describe the method used to reach a solution for a galaxy, given a set of models and a generalised parametrization. The construction of these models and choice of parameters is described in Sections 2.3 and 2.4.

We re-write the problem described by equation (4) in a simpler way

$$F_j = \sum_{\kappa=1}^{2S} c_\kappa A_{\kappa j}(Z_\kappa) \quad (10)$$

where $Z_\kappa = Z_a$ for $\kappa < S$ and $Z_\kappa = Z_b$ for $\kappa \geq S$. $\mathbf{A}$ is a $D \times 2S$ matrix composed of synthetic models at the corresponding metallicities, and $\mathbf{c} = (c_1, ..., c_{2S})$ is the solution vector, from which the x_α and g_α in equation (4) can be calculated. We can then calculate a linearly interpolated metallicity at age t_α

$$Z_\alpha = g_\alpha Z_a + (1 - g_\alpha) Z_b. \quad (11)$$

For every age t_α we aim to recover two parameters: x_α - the total mass formed at that age (within a period Δt_α) - and Z_α - a mass-weighted metallicity.

At this stage we are not concerned with our choice of t_α and Δt_α - although these are crucial and will be discussed later. For a given set of chosen parameters, we find $\mathbf{c}$, such that

$$\chi^2 = \frac{(F_j - \sum_\kappa c_\kappa A_{\kappa j})^2}{\sigma_j^2} \quad (12)$$

is minimised (where σ_j is the error in the measured flux bin F_j).

A linear problem with a least squares constraint has a simple analytic solution which, for constant $\sigma_\mathbf{j}$ (white-noise) is

$$\mathbf{c_{LS}} = (\mathbf{A^T} \cdot \mathbf{A})^{-\mathbf{1}} \cdot \mathbf{A^T} \cdot \mathbf{F} \quad (13)$$

In principle, any matrix inversion method, e.g. Singular Value Decomposition (SVD), can be used to solve (13). However, we would like to impose positivity constraints on the recovered solutions. Negative solutions are unphysical, but unfortunately common in a problem perturbed by noise.

2.2.1 BVLS and positivity

We use Bounded-Variable Least-Squares (BVLS) (Stark & Parker 1995) in order to solve (13). BVLS is an algorithm which solves linear problems whose variables are subject to upper and lower constraints. It uses an active set strategy to choose sets of free variables, and uses QR matrix decomposition to solve the unconstrained least-squares problem of each candidate set of free variables using (13):

$$\mathbf{c_{LS}} = (\mathbf{E^T} \cdot \mathbf{E})^{-\mathbf{1}} \cdot \mathbf{E^T} \cdot \mathbf{F} \quad (14)$$

where $\mathbf{E}$ is effectively composed of those columns of $\mathbf{A}$ for which c_k is unconstrained, and of zero vectors for those columns for which c_k is set to zero. BVLS is an extension of the Non-Negative Least Squares algorithm (Lawson & Hanson 1974), and they are both proven to converge in a finite number of iterations. Positivity is the only constraint in VESPA's solution.

BVLS and positivity have various advantages. Most obvious is the fact that we do away with negative solutions. In a non-constrained method (such as SVD) negative values are a response to the fact that the data is noisy. Similarly, we find that zero values returned by BVLS (in, for example, a synthetic galaxy with continuous star formation across all time) are also an artifact from noisy data. It should be kept in mind that, if the method is unbiased, this problem is solved by analysing a number of noisy realisations of the original problem - what we find is that the true values of the parameters we try to recover are consistent with the distributions yielded by this process. In this sense, not even a negative value presents a problem necessarily, as long as it is consistent with zero (or the correct solution). Given that we have found no bias when using BVLS, we feel it is an advantage to discard a priori solutions we know to be unphysical.

Another advantage to using BVLS is the fact that, by fixing some parameters to the lower boundary (zero, in this case), it effectively reduces the number of fitting parameters to the number of those which keeps unconstrained. Given the overall aims of VESPA, this has proven to be advantageous.

2.2.2 Noise

The inversion in equation (13) is often highly sensitive to noise, and care is needed when recovering solutions with matrix inversion methods. The fit in data-space will always improve as we increase the number of parameters, but these might not all provide meaningful information. We follow an analysis given in Ocvirk et al. (2006) in order to understand how much this affects our results, and to choose a suitable age parametrization for each galaxy. This is not an exact method, and it does not guarantee that the solutions we recover have no contribution from noise. However, we found that in most cases it provides a very useful guideline (see section 3.3, in particular Figure 11).

We refer the reader to the above paper for a full discussion, and we reproduce here the steps used in our analysis.

We use SVD to decompose the model matrix $\mathbf{E}$ as

$$\mathbf{E} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V^T} \quad (15)$$

where $\mathbf{U}$ is a $D \times 2S$ orthonormal matrix with singular data-vectors $\mathbf{u}_\kappa$ as columns, $\mathbf{V}$ is a $2S \times 2S$ orthonormal matrix with the singular solution-vectors $\mathbf{v}_\kappa$ as columns, and $\mathbf{W}$ is a $2S \times 2S$ diagonal matrix $\mathbf{W} = diag(w_1, ..., w_{2S})$ where w_κ are the matrix singular values in decreasing order. Replacing $\mathbf{E}$ by this decomposition in equation (13) gives

$$\mathbf{c}_{LS} = \mathbf{V} \cdot \mathbf{W}^{-1} \cdot \mathbf{U}^T \cdot \mathbf{F} = \sum_{\kappa=1}^{2S} \frac{\mathbf{u}_\kappa^T \cdot \mathbf{F}}{w_\kappa} \mathbf{v}_\kappa \quad (16)$$

The solution vector is a linear combination of the solution singular values, parametrized by the dot product between the data and the corresponding data singular vector, and divided by the k^{th} singular value. The data vector itself is a combination of the true underlying emitted flux and noise: $\mathbf{F} = \mathbf{F_{true}} + \mathbf{e}$. Equation (16) becomes

$$\mathbf{c}_{LS} = \sum_{\kappa=1}^{2S} \frac{\mathbf{u}_\kappa^T \cdot \mathbf{F}_{true}}{w_\kappa} \mathbf{v}_\kappa + \sum_{\kappa=1}^{2S} \frac{\mathbf{u}_\kappa^T \cdot \mathbf{e}}{w_\kappa} \mathbf{v}_\kappa \equiv \mathbf{c}_{true} + \mathbf{c}_e \quad (17)$$

where $\mathbf{c_{true}}$ is the solution vector to the noiseless problem and $\mathbf{c_e}$ is an unavoidable added term due to the presence of noise.

It is extremely informative to compare the amplitudes of the two terms in the sum (17), and to monitor their contributions to the solution vector with varying rank. In Figure 1 we plot $|\mathbf{u_\kappa^T} \cdot \mathbf{F}|$ and $\mathbf{u_\kappa^T} \cdot \mathbf{e}$ as a function of rank κ, for a synthetic spectrum with a SNR per pixel of 50 (at a resolution of 3Å) and an exponentially-decaying star formation history. We observe the behaviour described and discussed in Ocvirk et al. (2006). The combinations associated with the noise terms maintain a roughly constant power across all ranks, with a an average value of $\langle \mathbf{F} \rangle$ /SNR. The data terms, however, drop significantly with rank, and we can therefore identify two ranges: a noise-dominated κ-range, in which the noise contributions match or dominate the true data contributions, and a data-dominated range, where the contributions to the solution are largely data motivated. We call the transition rank κ_{crit}. Overall, high-κ ranks tend to dominate the solution, since the singular values w_κ decrease with κ. This only amplifies

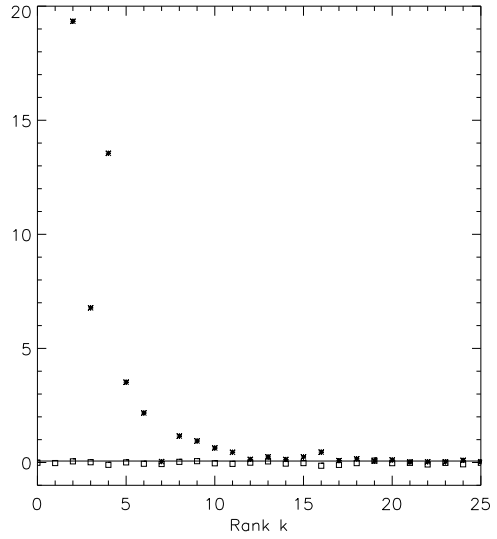

Figure 1. The behaviour of the singular values with matrix rank k. The stars are $|\mathbf{u}_\kappa^T \cdot \mathbf{F}|$ and the squares are $\mathbf{u}_\kappa^T \cdot \mathbf{e}$. The line is $\langle \mathbf{F} \rangle$ /SNR, which in this case has a value of approximately 0.06.

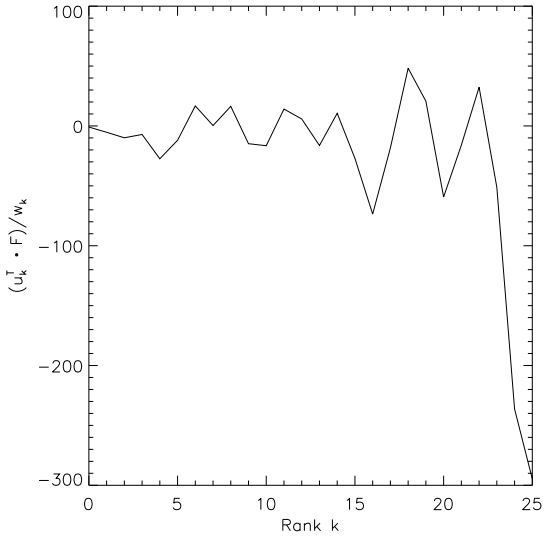

Figure 2. The coefficients in sum (16) as a function of rank κ. We see that the highest rank modes (corresponding to the smaller singular values) tend to contribute the most to the solution.

the problem by giving greater weight to noise-dominated terms in the sum (16). Figure 2 shows the contribution coming from each rank κ to the final solution - the coefficient $(\mathbf{u}_\kappa^T \cdot \mathbf{F})/w_\kappa$. We see this weight increases with rank.

Whereas this analysis gives us great insight into the problem, we do not in fact use the sum (16) to obtain $\mathbf{c}_{LS}$, for the reasons given in section 2.2.1.

For real data we are only able to calculate $\mathbf{u}_\kappa^T \cdot \mathbf{F}$ and estimate the noise level at $\langle \mathbf{F} \rangle$ /SNR and we use this information to estimate the number of non-zero parameters to recover from the data. Our aim is to a have a solution which is dominated by the signal, and not by the noise. We therefore want our number of non-zero recovered parameters to be less than or equal to κ_{crit}. Estimating where this transition happens is always a noisy process. In this paper we take the conservative approach of setting κ_{crit} to be the rank at which the perturbed singular values first cross the $\langle \mathbf{F} \rangle$ /SNR barrier. In the case of Figure 1 this happens at $\kappa_{crit} = 7$.

2.3 Choosing a galaxy parametrization

One of the advantages of VESPA is that it has the ability to choose the number of parameters to recover in any given galaxy. This is possible due to a time grid of varying resolutions, which VESPA can explore to find a solution. This section describes this grid and the criteria used to reach a final parametrization.

2.3.1 The grid

We work on a grid with a maximum resolution of 16 age bins, logarithmically spaced in lookback time from 0.02 up to 14 Gyr. The grid has three further resolution levels, where we split the age of the Universe in eight, four and finally two age bins, also logarithmically spaced in the same range.

The idea behind the multi-resolution grid is to start our search with a low number of parameters (in coarser resolution, so that the entirety of the age of Universe is covered), and then increase the resolution only where the data warrant it by splitting the bin with the highest flux contribution in two, and so on. In effect, we construct one such grid for each of the tabulated metallicities, Z_a and Z_b. We work with five metallicity values, $Z = [0.0004, 0.004, 0.008, 0.02, 0.05]$ which correspond to the metallicity resolution of the models used, where Z is the fraction of the mass of the star composed of metals ($Z_\odot = 0.02$). The construction of the models for each of the time bins is discussed in Section 2.4.

To each of the grids we can apply a dust extinction as explained in Section 2.1.1.

2.3.2 The search

We go through the following steps in order to reach a solution:

(i) We begin our search with three bins: two bins of width 4 and one bin of width 8 (oldest), where here we are measuring widths in units of high-resolution bins.

(ii) We calculate a solution using equation (10) for every possible combination of consecutive boundaries Z_a and Z_b, and we choose the one which gives the best value of reduced χ^2.

(iii) We calculate the number of perturbed singular values above the noise level, as described at the end of Section 2.2.2.

(iv) We find the bin which contributes the most to the total flux and we split it into two.

(v) We find a solution in the new parametrization, this time by trying out all possible combinations of Z_a and Z_b for the newly split bins only, and fixing the metallicity boundaries of the remainder bins to the boundaries obtained in the previous solution. If a bin had no stars in the previous iteration, we set $Z_a = 0.0004$ and $Z_b = 0.05$.

(vi) We return to (iii) and we proceed until we have reached the maximum resolution in all populated bins.

(vii) We look backwards in our sequence of solutions for the last instance with a number of non-zero recovered parameters equal to or less than κ_{crit} as calculated in (iii) and take this as our best solution.

We illustrate this sequence in Figure 3, where we show the evolution of the search in a synthetic galaxy composed of two stellar bursts of equal star formation rates - one young and one old. VESPA first splits the components which contribute the most to the total flux. In this case this is the young burst which can be seen in the first bin. Even though VESPA always resolves bins with any mass to the possible highest resolution, it then searches for the latest solution which has passed the SVD criterion explained in Section 2.2.2. In this case, this corresponds to the fifth from the top solution. VESPA chooses this solution in favour of the following ones due to the number of perturbed singular values above the solid line (right panel). In this case, the solution chosen by VESPA is a better fit in parameters space (note the logarithmic scale in the y-axis - the following solution put the vast majority of the mass in the wrong bin). We observed this type of improvement in the majority of all cases studied (see Figure 11).

2.3.3 The final solution

Our final solution comes in a parametrization such that the total number of non-zero recovered parameters is less than or equal to the number of perturbed singular values above the estimated noise level.

The above sequence is performed for each of several combinations of $\tau_V^{BC}, \tau_V^{ISM}$, and we choose the attenuation which provides the best fit.

For each galaxy we recover N star formation masses, with an associated metallicity, where N is the total number of bins, and a maximum of two dust parameters.

2.4 The models

The backbone to our grid of models is the BC03 set of synthetic SSP models (Bruzual & Charlot 2003), with a Chabrier initial mass function (Chabrier 2003) and Padova 1994 evolutionary tracks (Alongi et al. 1993; Bressan et al. 1993; Fagotto et al. 1994a,b; Girardi et al. 1996) . Although any set of stellar population models can be used, these provide a detailed spectral evolution of stellar populations over a suitable range of wavelength, ages and metallicities: $S(\lambda, t, Z)$. The models have been normalised to one solar mass at the age $t = 0$.

2.4.1 High-resolution age bins

At our highest resolution we work with 16 age bins, equally spaced in a logarithmic time scale between now and the age of the Universe. In each bin, we assume a constant star formation rate

$$f_\alpha^{HR}(\lambda, Z) = \psi \int_{\Delta t_\alpha} S(\lambda, t, Z) dt \tag{18}$$

with $\psi = 1/\Delta t_\alpha$.

2.4.2 Low-resolution age bins

As described in Section 2.3.1, we work on a grid of different resolution time bins and we construct the low resolution bins using the high resolution bins described in Section 2.4.1. We do not assume a constant star formation rate in this case, as in wider bins the light from the younger components would largely dominate over the contribution from the older ones. Instead, we use a decaying star formation history, such that the light contributions from all the components are comparable. Recall equation (1)

$$f_\alpha^{LR}(\lambda, Z) = \int_{\Delta t_\alpha} \psi(t) S(\lambda, t, Z) dt, \tag{19}$$

which we approximate to

$$f_\beta^{LR}(\lambda, Z) = \frac{\sum_{\alpha \in \beta} f_\alpha^{HR}(\lambda, Z) \psi_\alpha \Delta t_\alpha}{\sum_{\beta \in \alpha} \psi_\alpha \Delta t_\alpha} \tag{20}$$

where low resolution bin β incorporates the high resolution bins $\alpha \in \beta$, and we set

$$\psi_\alpha \Delta t_\alpha = \frac{1}{\int_\lambda f_\alpha^{HR}(\lambda, Z) d\lambda}. \tag{21}$$

Depending on the galaxy, the final solution obtained with the sequence detailed in Section 2.3.2 can be described in terms of low-resolution age bins. In this case we should interpret the recovered mass as the total mass formed during the period implied by the width of the bin, but we cannot make any conclusions as to when in the bin the mass was formed. Similarly, the recovered metallicity for the bin should be interpreted as a mass-weighted metallicity for the total mass formed in the bin.

2.5 Errors

The quality of our fits and of our solutions is affected by the noise in the data, the noise in the models, and the parametrization we choose (which does not reflect the complete physical scenario within a galaxy). We aim to apply VESPA firstly to SDSS galaxies, which typically have a SNR ≈ 20 per resolution element of 3Å, which puts us in a regime where the main limitations come from the noise in the data.

To estimate how much noise affects our recovered solutions we take a rather empirical approach. For each recovered solution we create n_{error} random noisy realisations and we apply VESPA to each of these spectra. We re-bin each recovered solution in the parametrization of the solution we want to analyse and estimate the covariance matrices

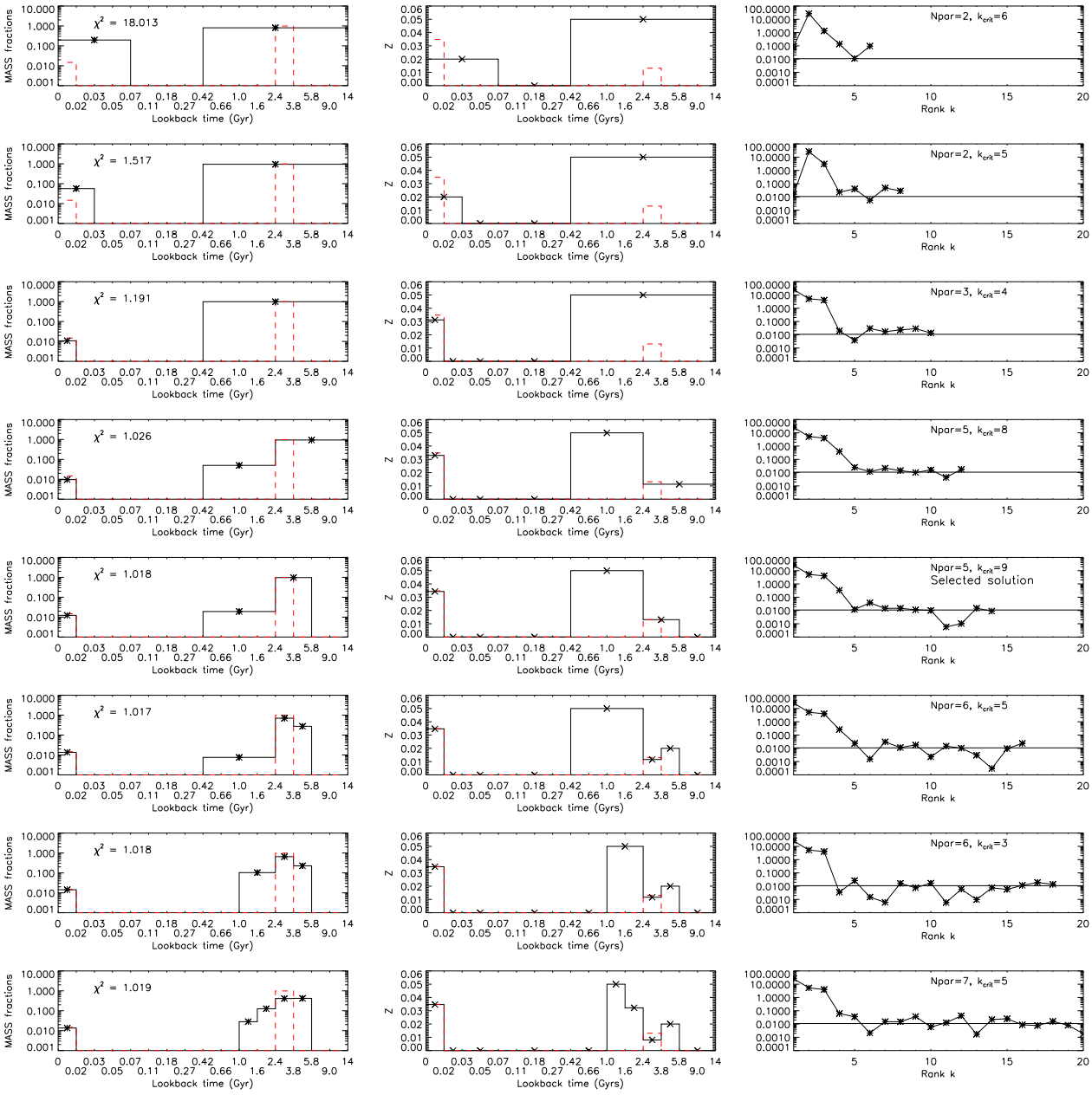

Figure 3. The evolution of the fit, as VESPA searches for a solution. Sequence should be read from top to bottom. Each line shows a stage in the sequence: the left panel shows the input star formation history in the dashed line (red on the online version), and the recovered mass fractions on the solid line (black on the online version) for a given parametrization ; the middle panel shows the input metallicities in the dashed line (red on the online version), and the recovered metallicities on the solid line (black on the online version); the right panel shows the absolute value of the perturbed singular values $|\mathbf{u}_\kappa \cdot \mathbf{F}|$ (stars and solid line) and the estimated noise level $\langle \mathbf{F} \rangle$ /SNR. In this panel we also show the value of κ_{crit} and the number of non-zero elements of $\mathbf{c_{LS}}$ in each iteration. The chosen solution is the **fifth from the top**, and indicated accordingly. This galaxy consists of 2 burst events of equal star formation rate - a very young and an old burst. It was modelled with a resolution of 3Å and a signal-to-noise ratio per pixel of 50. We see the recovery is good but not perfect - there is a 1 per cent leakage from the older population - but better than the following solutions, where this bin is split. See text in Section 2.3.2 for more details.

$$C(x)_{\alpha\beta} = \langle (x_\alpha - \bar{x}_\alpha)(x_\beta - \bar{x}_\beta) \rangle \quad (22)$$

$$C(Z)_{\alpha\beta} = \left\langle (Z_\alpha - \bar{Z}_\alpha)(Z_\beta - \bar{Z}_\beta) \right\rangle . \quad (23)$$

All the plots in Sections 3 and 4 show error bars derived from $C_{\alpha\alpha}^{1/2}$, although it is worth keeping in mind that these are typically highly correlated.

2.6 Timings

A basic run of VESPA (which consists of roughly 5 runs down the sequence detailed in Section 2.3.2, one for each value of dust extinction) takes about 5 seconds. If accurate error estimations are needed per galaxy, this will add another one or two minutes to the timing, depending on how

accurately one would like to estimate the covariance matrices, and depending on the number of data points. With $n_{error} = 10$, a typical SDSS galaxy takes around one minute to analyse.

3 TESTS ON SIMULATED DATA

We tested VESPA on a variety of synthetic spectra, in order to understand its capabilities and limitations. In particular, we tried to understand the effect of three factors in the quality of our solutions: the input star formation history, the noise in the data, and the wavelength coverage of the spectrum. We have also looked at the effects of dust extinction. Throughout we have modelled our galaxies in a resolution of 3Å.

Even though we are aware that showing individual examples of VESPA's results from synthetic spectra can be extraordinarily unrepresentative, we feel obliged to show a few for illustration purposes. We will show a typical result for most of the cases we present, but we also define some measurements of success, so that the overall performance of VESPA can be tracked as we vary any factors. We define

$$G_x = \sum_\alpha \left| \frac{x_\alpha - x_\alpha^I}{x_\alpha^I} \right| \omega_\alpha \tag{24}$$

and

$$G_Z = \sum_\alpha \left| \frac{Z_\alpha - Z_\alpha^I}{Z_\alpha^I} \right| \omega_\alpha \tag{25}$$

where x_α^I and Z_α^I are the total mass and correspondent metallicity in bin α (re-binned to match the solution's parametrization if necessary), and ω_α is the flux contribution of population of age t_α. G_x and G_Z are a flux-weighted average of the total absolute fractional errors in the solution, and give an indication of how well VESPA recovers the most significant parameters. A perfect solution gives $G_x = G_Z = 0$. It is also worth noting that this statistic does not take into account the associated error with each recovered parameter - deviations from the true solution are usually expected given the estimated covariance matrices. We will also show how these factors affect the recovered total mass for a galaxy. In all cases we have re-normalised the total masses such that total input mass for each galaxy is 1.

3.1 Star formation histories

We present here some results for synthetic spectra with two different star formation histories. All of the spectra in this section were synthesised with a SNR per pixel of 50, and we initially fit the very wide wavelength range $\lambda \in [1000, 9500]$Å.

We choose two very difference cases: firstly a star formation history of dual bursts, with a large random variety of burst age separations and metallicities (where we set the star formation rate to be 10 solar masses per Gyr in all bursts). Secondly, we chose a SFH with an exponentially decaying star formation rate: SFR $\propto \exp(\gamma t_\alpha)$, where t_α is the age of the bin in lookback time in Gyr. Here we show results for $\gamma = 0.3$ Gyr^{-1}. Rather than being physically motivated, our choice of γ reflects a SFH which is not too steep as to essentially mimic a single old burst, but which is also not completely dominated by recent star formation. In all cases the metallicity in each bin is randomly set. Figure 4 shows a typical example from each type.

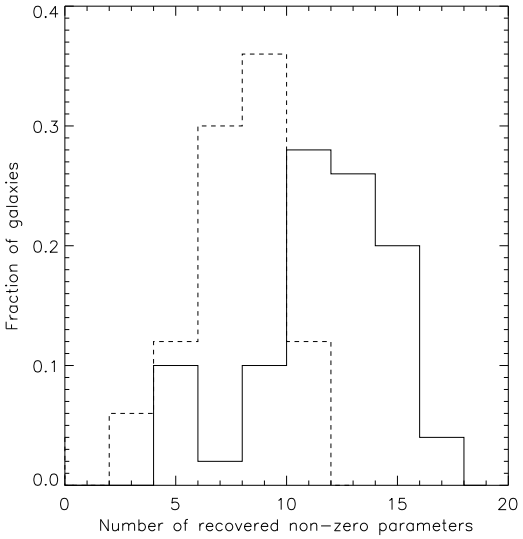

Figure 8. The recovered number of non-zero parameters in 50 galaxies with an exponentially decaying star formation history, using different wavelength ranges: $\lambda \in [1000, 9500]$ Å(solid line) and $\lambda \in [3200, 9500]$ Å(dashed line). Please note that these correspond to the *total* number of non-zero components in the solution vector $\mathbf{c}_\kappa$ and not to the number of recovered stellar populations.

Figure 5 shows the distribution of G_x, G_Z and of the recovered total masses for a sample of 50 galaxies. We see differences between the two cases. Firstly, in dual bursts galaxies, we seem to do better in recovering data from significant individual bins, but worse in overall mass. This reflects the fact that G_x is dominated by the fractional errors in the most significant bins, but the total mass can be affected by small flux contributions in old bins which can have large masses. On the other hand, with an exponentially decaying star formation rate, we do worse overall (although this is mainly a reflection that more bins have significant contributions to the flux) but we recover the total mass of the galaxy exceptionally well.

3.2 Wavelength range

Wavelength range is an important factor in this sort of analysis, as different parts of the spectrum will help to break different degeneracies. Since we are primarily interested in SDSS galaxies, we have studied how well VESPA does in the more realistic wavelength range of $\lambda \in [3200, 9500]$ Å.

Figure 6 shows the results for the same galaxies shown in Figure 4, obtained with the new wavelength range. In these

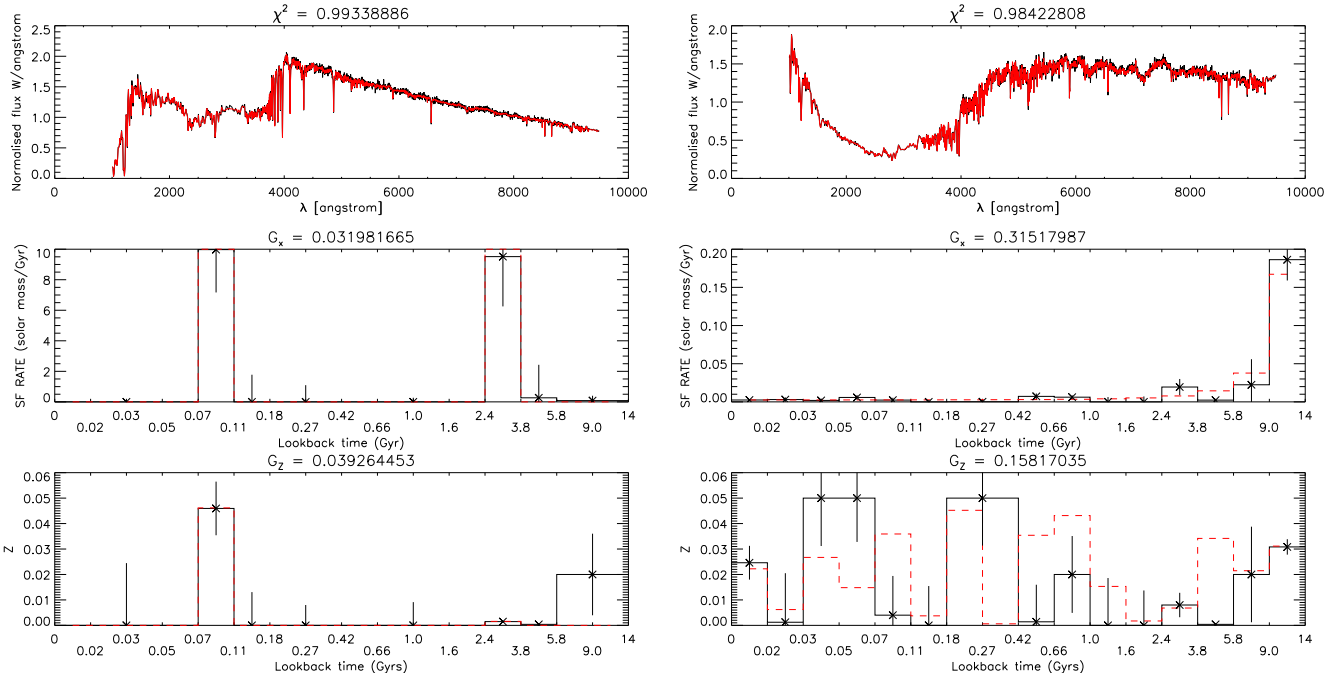

Figure 4. Two examples of VESPA's analysis on synthetic galaxies. The top panels show the original spectrum in the dark line (black in the online version) and fitted spectrum in the lighter line (red in the online version). The middle panels show the input (dashed, red) and the recovered (solid, black) star formation rates and the bottom panel shows the input (dashed, red) and recovered (solid, black) metallicities per bin. Note that even though many of the recovered metallicities are wrong, these tend to correspond to bins with very little star formation, and are therefore virtually unconstrained.

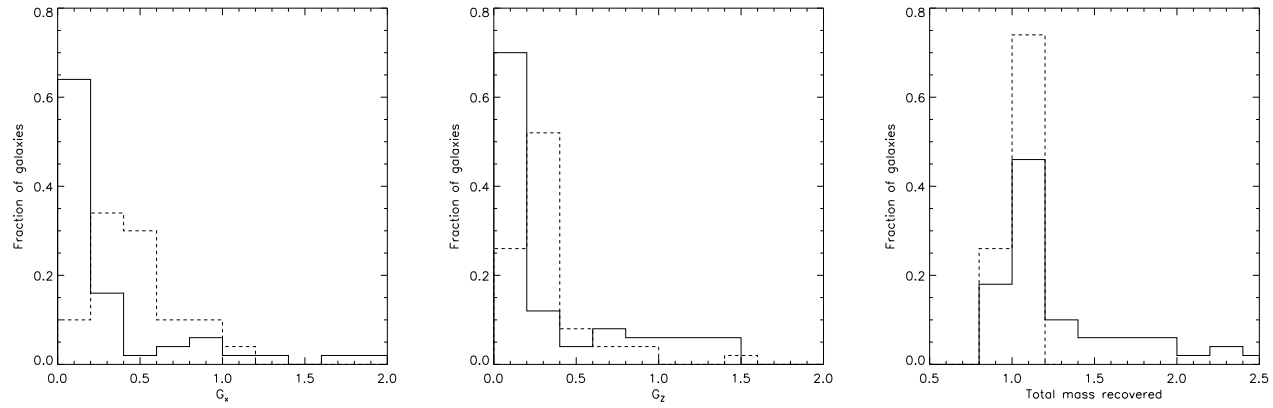

Figure 5. The distribution of G_x, G_Z and total mass recovered for 50 galaxies with a SNR per pixel of 50. Solid lines correspond to dual burst and dashed lines to exponentially decaying ones. See text in Section 3.1 for details.

particular cases, we notice a more pronounced difference in the dual bursts galaxy, but looking at a more substantial sample of galaxies shows that this is not generally the case. Figure 7 shows G_x, G_Z and total mass recovered for 50 exponentially decaying star formation history galaxies, with a signal to noise ratio of 50 and the two different wavelength ranges. We do not see a largely significant change in both cases, and we observe a less significant difference in the dual bursts galaxies (not plotted here).

We find it instructive to keep track of how many parameters we recover in total, as we change any factors. Figure 8 shows an histogram of the total number of non-zero parameters we recovered from our sample galaxies with exponentially-decaying star formation histories and both wavelength ranges. Note that these are the components of the solution vector $\mathbf{c}_\kappa$ which are non-zero - they do not represent a number of recovered stellar populations. In this case there is a clear decrease in the number of recovered parameters, suggesting a wider wavelength range is a useful way to increase resolution in parameter space.

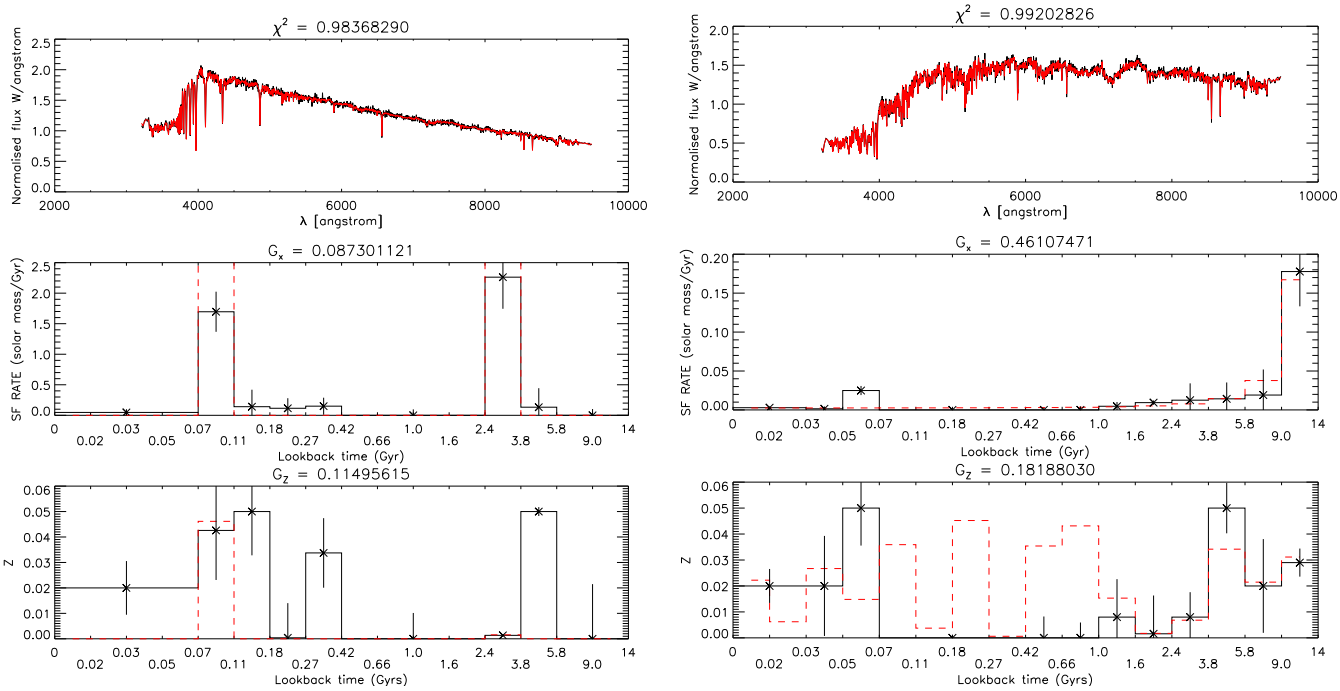

Figure 6. Same galaxies as in Figure 4, but results obtained by using a smaller wavelength range. The goodness-of-fit in data space is still excellent, but it becomes more difficult to break certain degeneracies.

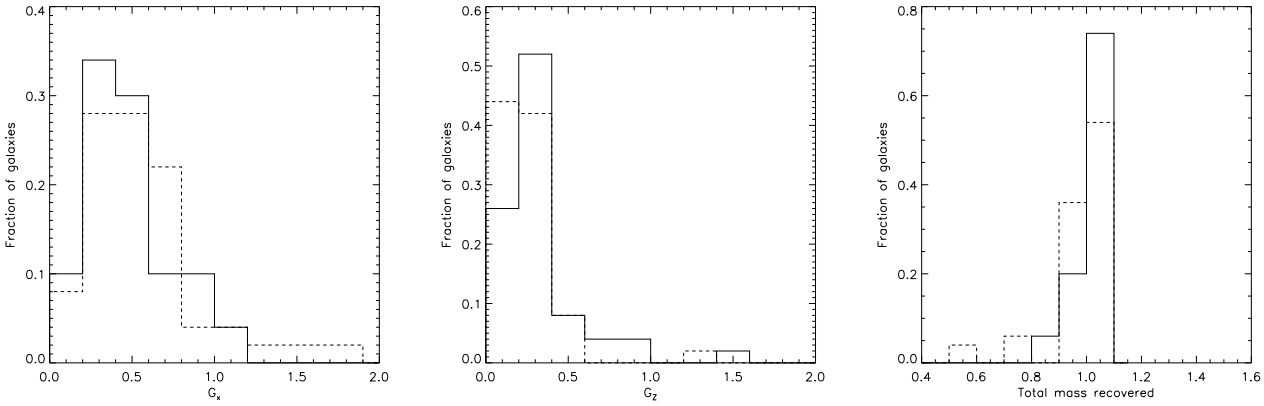

Figure 7. The distribution of G_x, G_Z and total mass recovered for 50 galaxies with a SNR per pixel of 50 and two different wavelength coverage. Solid line corresponds to $\lambda \in [1000, 9500]\mathring{A}$ and dashed line to $\lambda \in [3200, 9500]$ Å.

3.3 Noise

It is of interest to vary the signal-to-noise ratio in the synthetic spectra. We have repeated the studies detailed in the two previous sections with varying values of noise, and we investigate how this affects both the quality of the solutions and their resolution in parameter space.

Figure 10 shows how the recovered number of parameters changed by increasing the noise in the galaxies with an exponentially decaying star formation rate and wide wavelength range. In this case the increase in the noise leads to a significant reduction of the number of parameters recovered for each galaxy. This behaviour is equally clear for different star formation histories and different wavelength coverage, and is directly caused by the stopping criterion defined in Section 3.3.

The quality of the solutions is also affected by this increase in noise, as can be seen in figure 9, where we have plotted G_x, G_Z and the total recovered mass for two different values of SNR. The quality of the solutions decreases with the higher noise levels, as is to be expected. However, a more interesting question to ask is whether this decrease in the quality of the solutions would indeed be more pronounced without the SVD stopping criterion. Figure 11 shows a comparison between G_x obtained as we have described and obtained without any stopping mechanism (so letting our search go to the highest possible resolution and taking the final solution) for 50 galaxies with an exponentially decaying star formation history and a signal-to-noise ratio of 20. The re-

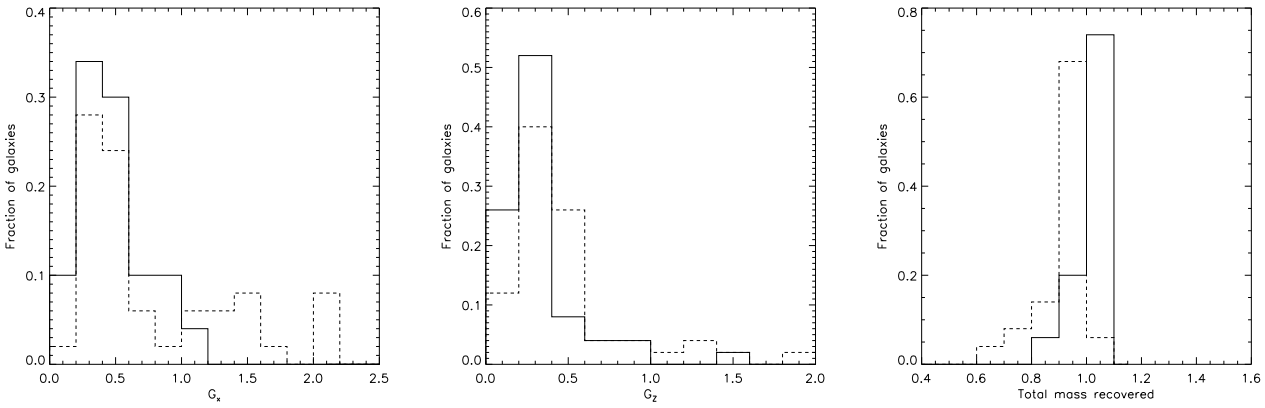

Figure 9. The distribution of G_x, G_Z and total mass recovered for 50 galaxies with an exponential decaying star formation history and different signal-to-noise ratios. Solid lines correspond to SNR = 50 and dashed lines to SNR =20. See text in Section 3.3 for details.

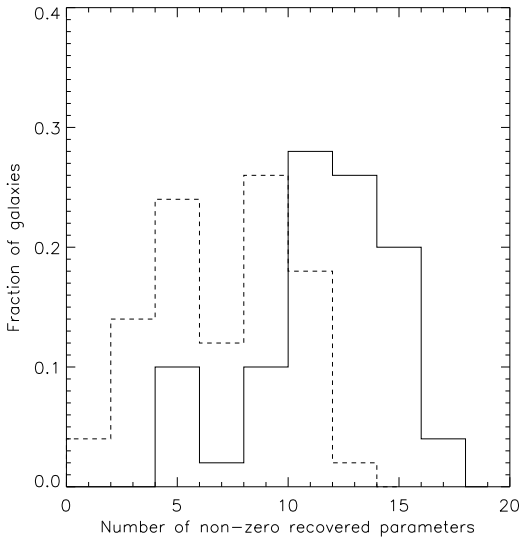

Figure 10. The recovered number of non-zero parameters as we change the noise in the data from 50 (solid line) to 20 (dashed line), in a sample of galaxies with an exponentially decaying star formation rate. Please note that these correspond to the *total* number of non-zero components in the solution vector $\mathbf{c}_\kappa$ and not to the number of recovered stellar populations.

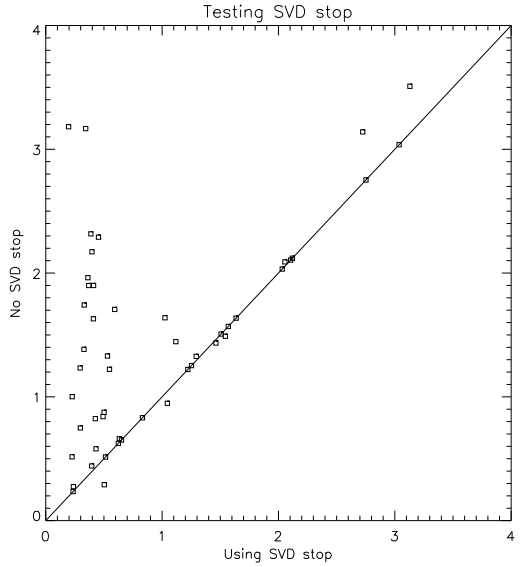

Figure 11. Testing the SVD stopping criterion. Plots show goodness of fit G_x for the solution of 50 galaxies obtained with and without the SVD stopping criterion. We see that by recovering only as many parameters as the data warrants gives improved parameter estimation in almost all cases, and a striking improvement in many.

sults show clearly that there is a significant advantage in using the SVD stopping criterion. Naturally, the goodness of fit in data space is consistently better as we increase the number of parameters but this improvement is illusory - the parameter recovery is worse. This is exactly the expected behaviour - we choose to sacrifice resolution in parameter space in favour of a more robust solution - even though naively one could think a lower χ^2 solution would indicate a better solution. The significance of this improvement changes with the amount of noise and wavelength range of the data (and to a lesser extent with type of star formation history) but we observed an improvement in all cases we have studied.

As expected, further decreasing the signal-to-noise ratio leads to a further degradation of the recovered solutions. This is accompanied by a suitable increase in the error bars and correlation matrices, but in cases of a SNR$\approx$ 10 and less it becomes very difficult to recover any meaningful information from individual spectra.

3.4 Dust

In this section we use simulated galaxies to study the effect of dust in our solutions. As explained in Section 2.1.1, due to the non-linear nature of the problem, we cannot include dust as one of the free parameters analysed by

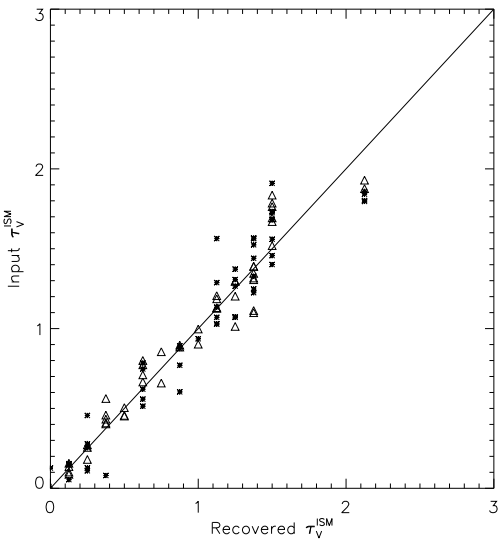

Figure 12. Testing the recovery of τ_V^{ISM} for 50 galaxies with a exponentially-decaying star formation history (triangles) and 50 galaxies formed with a random combination of dual bursts (stars). The input values are randomly chosen and continuously distributed between 0 and 2. The recovered values are chosen from a tabulated grid between 0 and 4.

SVD. Instead, we fit for a maximum of two dust parameters using a brute-force approach which aims to minimise χ^2 in data-space by trying out a series of values for τ_V^{ISM} and τ_V^{BC}.

For each galaxy we assign random values of $\tau_V^{ISM} \in [0, 2]$ and $\tau_V^{BC} \in [1, 2]$ and we are interested in how well we recover these parameters and any possible degeneracies.

Figure 12 shows the input and recovered values for τ_V^{ISM} for galaxies with a signal-to-noise ratio of 50, and which were analysed using the wavelength range $\lambda \in [3200, 9500]$Å. We show results for two different cases of star formation history: 50 galaxies with an exponentially-decaying SFR and 50 galaxies formed by dual-bursts. We observe a good recovery of τ_V^{ISM} in both cases, especially at low optical depths.

However, we mostly observe a poor recovery of τ_V^{BC}, especially at high optical depths. This is unsurprisingly flagging up a certain level of degeneracy between mass and degree of extinction, which gets worse as the optical depth increases. Essentially, it becomes difficult to distinguish between a highly obscured massive population and a less massive population surrounded by less dust. It is worth keeping in mind that young populations are affected by both dust components simultaneously, and generally, even though the recovery of the second dust parameter may not be accurate, it allows for a better estimation of the dominant dust component.

This can be tested by simulating galaxies on a two-component dust model and by analysing them using both a single component model, and a two-component model. E.g., when using the more sophisticated model, we noted that the mean error on τ_V^{ISM} on a subsample of dual-burst galaxies (synthesised as explained in section 3.1, but chosen to have young star formation) was reduced from 35 to 28 per cent. This simple test also revealed that we are less likely to underestimate the mass of young populations by allowing an extra dust component, but that we are also introducing an extra degeneracy, especially so in the case of faint young populations. However, we feel that the two-parameter dust model brings more advantages than disadvantages, with the caveat being that dusty young populations can be poorly constrained. In any case, we note that each galaxy is always analysed with a one-parameter model before being potentially analysed with a two-parameter model, and both solutions are kept and always available for analysis.

Finally, our test also partly justifies our choice to first run a single dust component model and only apply a two-component model if we detect stars in the first two bins - we find that although a one-component model might underestimate the amount of young stars, it does not fail to detect them. We repeated a similar test in real data, by analysing the same sample with and one- and a two-parameter dust model. We found similar results, with a one-parameter model failing to yield star formation in young bins only around 1 per-cent of the time (compared to the two-parameter model), and only in cases where the contribution of the light from the young populations was very small (of the order of 1 to 2 per cent).

4 RESULTS

In this section we present some results obtained by applying VESPA to galaxies in the SDSS. Our aim is to analyse these galaxies, and to produce and publish a catalogue of robust star formation histories, from which a wealth of information can then be derived. We leave this for another publication, but we present here results from a sub-sample of galaxies, which we used to test VESPA in a variety of ways.

4.1 Handling SDSS data

Prior to any analysis, we processed the SDSS spectroscopic data, so as to accomplish the desired spectral resolution and mask out any wanted signal.

The SDSS data-files supply a mask vector, which flags any potential problems with the measured signal on a pixel-by-pixel basis. We use this mask to remove any unwanted regions and emission lines. In practical terms, we ignore any pixel for which the provided mask value is not zero.

The BC03 synthetic models produce outputs at a resolution of 3Å, which we convolve with a Gaussian velocity dispersion curve with a stellar velocity $\sigma_V = 170 \mathrm{km} s^{-1}$, this being a typical value for SDSS galaxies. We take the models' tabulated wavelength values as a fixed grid and re-bin the SDSS data into this grid, using an inverse-variance weighted average. We compute the new error vector accordingly. Note that the number of SDSS data points averaged into any new

bin is not constant, and that the re-binning process is done after we have masked out any unwanted pixels. Additionally to the lines yielded by the mask vector, we mask out the following emission line regions in every spectrum's rest-frame wavelength range: [5885-5900, 6702-6732, 6716-6746, 6548-6578, 6535-6565, 6569-6599, 4944-4974, 4992-5022, 4846-4876, 4325-4355, 4087-4117, 3711-3741, 7800-11000] Å.

These re-binned data- and noise-vectors are essentially the ones we use in our analysis. However, since the linear algebra assumes white-noise, we pre-whiten the data and construct a new flux vector $F'_j = F_j/\sigma_j$, which has unit variance, $\sigma'_j = 1, \forall j$, and a new model matrix $A'_{ij} = A_{ij}/\sigma_j$.

4.2 Duplicate galaxies

There are a number of galaxies in the SDSS database which have been observed more than once, for a variety of reasons. This provides an opportunity to check how variations in observation-dependent corrections affect the results obtained by VESPA.

We have used a subset of the sample of duplicate objects in Brinchmann et al. (2004)[1] to create two sets of oservations for 2000 galaxies, which we named list A and list B. We are interested in seeing how the errors we estimate for our results compare to errors introduced by intrinsic variations caused by changing the observation conditions (such as quality of the spectra, placement of the fibre, sky subtraction or spectrophometric calibrations).

Figure 13 shows the average star formation fraction as a function of lookback time for both sets of observation. The error bars showed are errors on the mean. We see no signs of being dominated by systematics when estimating the star formation fraction of a sample of galaxies.

Figure 14 shows the total stellar mass obtained for a set of 500 galaxies in both observations (details of how we estimate the total stellar mass of a galaxy are included in section 4.4). The error bars are obtained directly from the estimated covariance matrix $C(x)$ (equation 22). Even though most of the galaxy duplicates produce mass estimates in agreement with each other given the error estimates, a minority does not. Upon inspection, these galaxies show significant differences in their continuum, but after further investigation it remains unclear what motivates such a difference. The simplest explanation is that the spectrophotometric calibration differs significantly between both observations, and that might have been the reason the plate or object was re-observed. Whatever the reason however, the clear conclusion is that stellar mass estimates are highly sensitive to changes in the spectrum continuum, and the errors we estimate from the covariance matrix alone might be too small.

We did not find any signs of a systematic bias in any of the analysis we carried out.

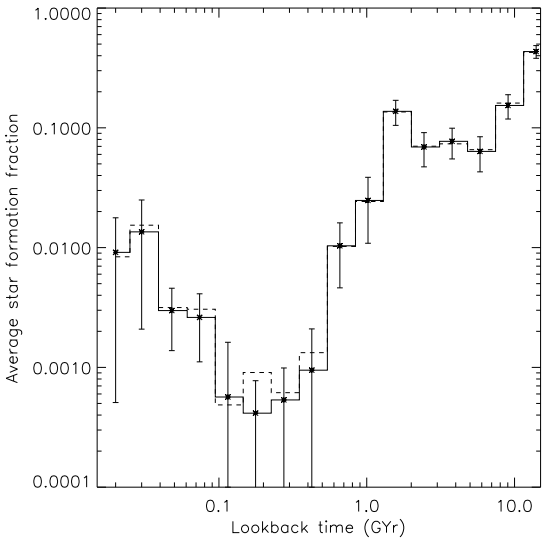

Figure 13. Average star formation fraction as a function of lookback time for the 2000 galaxies in list A (solid line) and list B (dashed line). The error bars shown are the errors bars on the mean for each age bin. We show only the errors from list A to avoid cluttering - the errors from list B are of similar amplitude.

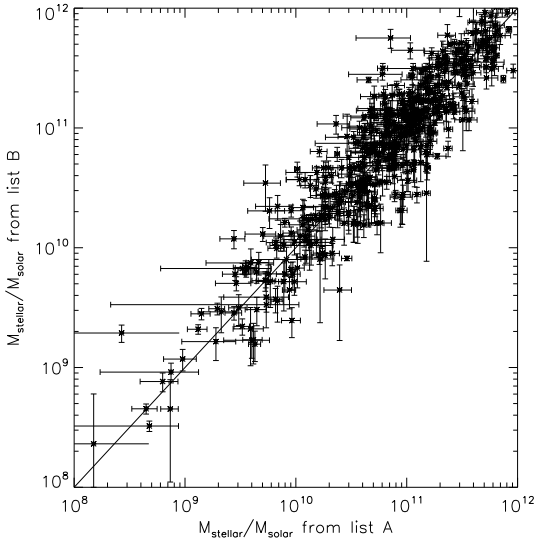

Figure 14. Total stellar mass recovered for two sets of observations of 500 galaxies in the main galaxy sample. The error bars are calculated from $C(x)$.

4.3 Real fits

In this section we discuss the quality of the fits to SDSS galaxies obtained with VESPA.

As explained in Section 2, VESPA finds the best fit solution in a χ^2 sense for a given parametrisation, which is self-

[1] Available at http://www.mpa-garching.mpg.de/SDSS/

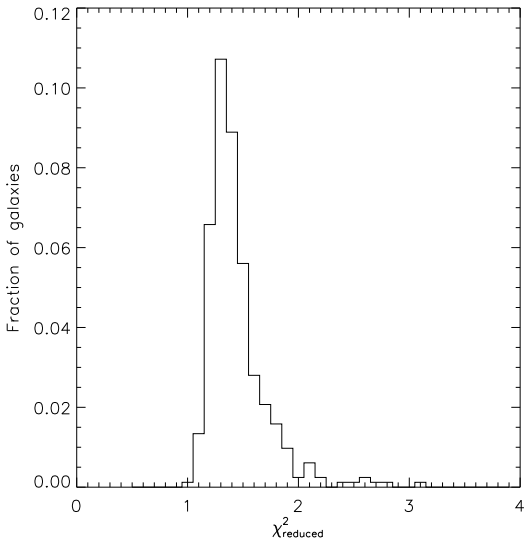

Figure 15. The distribution of reduced values of χ^2 for a sample of 360 galaxies analysed by VESPA.

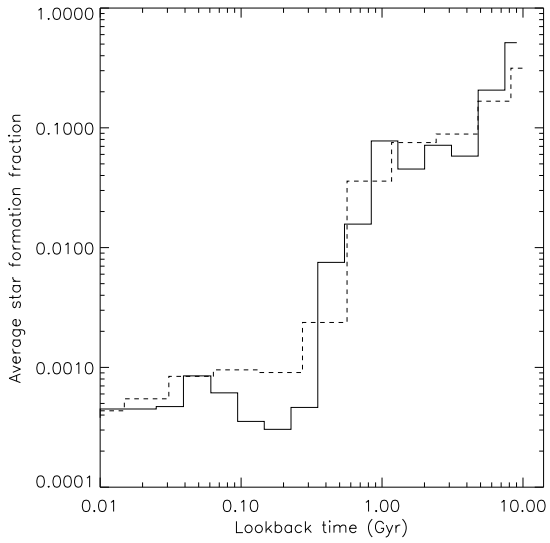

Figure 17. The recovered average star formation history for the 821 galaxies as recovered by VESPA (solid line) and MOPED (dashed line). Both were initially normalised such that the sum over all bins is 1, and the MOPED line was then adjusted by 11/16 to account for the different number of bins used in each method, to facilitate direct comparison.

regulated in order to not allow an excessive number of fitting parameters. We have shown that this self-regularization gives a better solution in parameter space (Figure 11), despite often not allowing the parametrization which would yield the best fit in data space (Figure 3). However, our aim is still to find a solution which gives a good fit to the real spectrum. Figure 15 shows the 1-point distribution of reduced values of reduced χ^2 for 1 plate of galaxies. This distribution peaks at around $\chi^2_{reduced} = 1.3$, and figure 16 shows a fit to one of the galaxies with a typical value of goodness of fit.

It is worth noting that the majority of the fits which are most pleasing to the eye, correspond to the ones with a high signal to noise ratio and high value of reduced χ^2. One would expect the best fits to come from the galaxies with the best signal. However, we believe the fact that they do not is not a limitation of the method, but a limitation of the modelling. There are a number of reasons why VESPA would be unable to produce very good fits to the SDSS data. One is the adoption of a single velocity dispersion (170 kms^{-1}) which could easily be improved upon at the expense of CPU time. However, the dominant reason is likely to be lack of accuracy in stellar and dust modelling - whereas BC03 models can and do reproduce a lot of the observed features, it is also well known that this sucess is limited as there are certain spectral features not yet accurately modelled, or even modelled at all. There are similar deficiencies in dust models and dust extinction curves. The effect of the choice of modelling should not be overlooked, and we refer the reader to a discussion in Section 4.5 of Panter et al. (2006)), where these issues are discussed.

4.4 VESPA and MOPED

In this Section we take the opportunity to compare the results from VESPA and MOPED, obtained from the same sample of galaxies. The VESPA solutions used here are obtained with a one-parameter dust model, to allow a more fair comparison between the two methods. Both methods make similar assumptions regarding stellar models, but MOPED uses an LMC (Gordon et al. 2003) dust extinction curve, and single screen modelling for all optical depths.

Our sample consist of two plates from the SDSS DR3 (Abazajian et al. 2005) (plates 0288 and 0444), from which we analyse a total of 821 galaxies. We are mainly interested in comparing the results in a global sense. MOPED in its standard configuration attempts to recover 23 parameters (11 star formation fractions, 11 metallicities and 1 dust parameter), so we might expect considerable degeneracies. Indeed, in the past the authors of MOPED have cautioned against using it to interpret individual galaxy spectra too precisely. We have observed degeneracies between adjacent bins in MOPED, but on the other hand a typical MOPED solution has many star formation fractions which are essentially zero, so the number of significant contributions is always much less than 23.

Figure 17 shows the recovered average star formation history for the 821 galaxies using both methods. In the case of VESPA, solutions parametrized by low-resolution bins had to be re-parametrized in high-resolution bins, so that a common grid across all galaxies could be used. This was done using the weights given by (21). The lines show a remarkably good agreement between the two methods.

Having recovered a star formation history for each galaxy, one can then estimate the stellar mass of a galaxy. We calculated this quantity for all galaxies using the solutions from both methods, and with similar assumptions regarding cos-

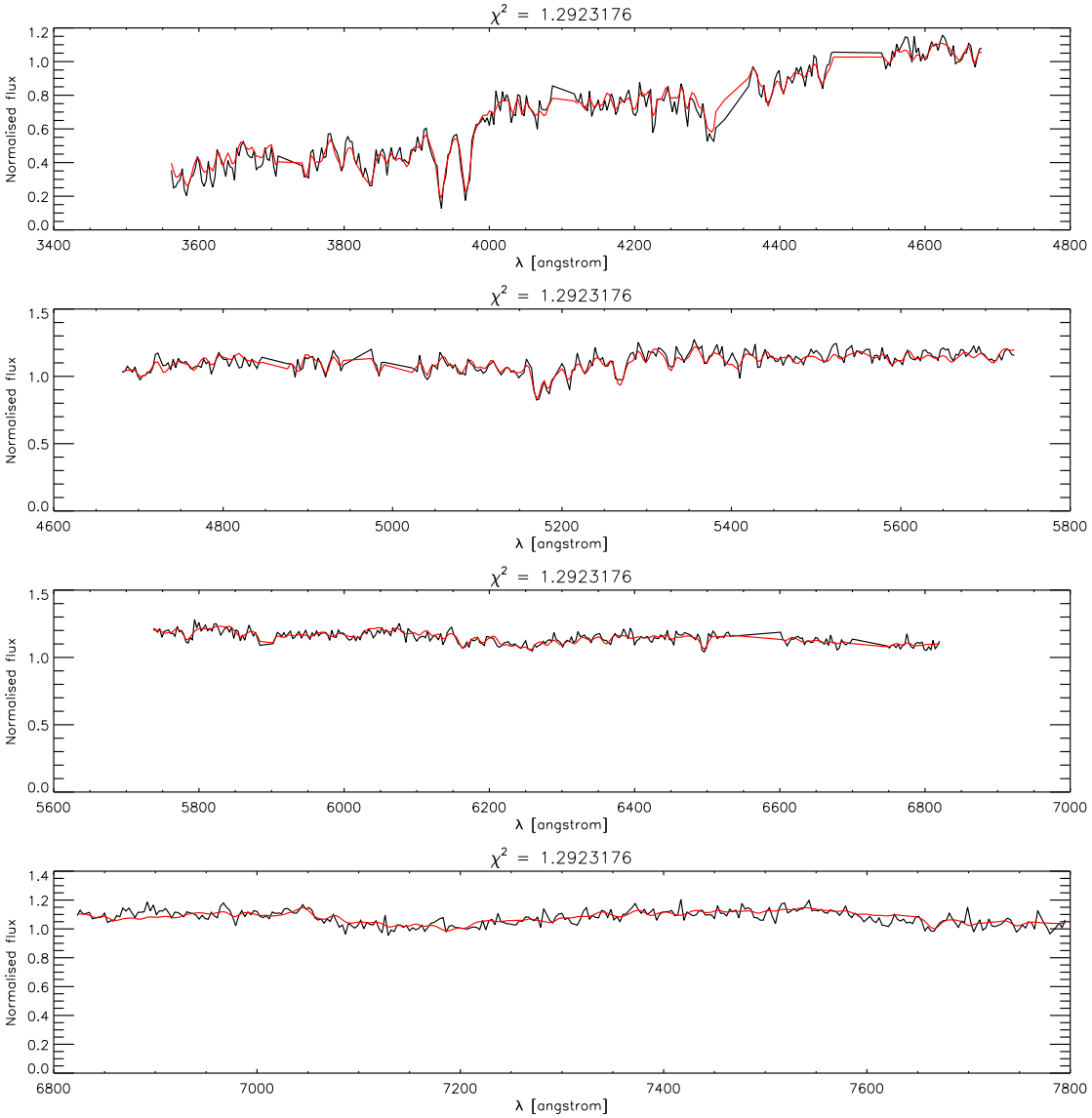

Figure 16. Typical fit to a galaxy from the SDSS. The dark line is the real data (arbitrary normalisation), and the lighter line (red on the online version) is VESPA's fit to the data.

mological parameters and fibre-size corrections. Explicitly, we have done the following:

(i) We converted from flux to luminosity assuming the set of cosmological parameters given by Spergel et al. (2003).

(ii) We recovered the initial mass in each age bin using each method.

(iii) We calculated the remaining present-day mass for each population after recycling processes. This information is supplied by the synthetic stellar models, as a function of age and metallicity.

(iv) We summed this across all bins to calculate the total stellar present-day mass in the fibre aperture, M.

(v) We corrected for the aperture size by scaling up the mass to $M_{stellar}$ using the petrosian and fibre magnitudes in the z-band, $M_p(z)$ and $M_f(z)$, with: $M_{stellar} = M \times 10^{0.4[M_p(z)-M_f(z)]}$

Figure 18 shows the recovered galaxy masses as recovered from MOPED and from VESPA. We see considerable agreement between VESPA and MOPED. Over 75 per cent of galaxies have $0.5 \leq M_{VESPA}/M_{MOPED} \leq 1.5$. There is a tail of around 10 per cent of galaxies where VESPA recovers 2 to 4 times the mass recovered by MOPED. The main reason for this difference is in the dust model used - we find a correlation between dust extinction and the ratio of the two mass estimates. This again reflects the fact that total stellar mass estimates are highly sensitive to changes in the spectrum continuum (see also section 4.2).

Our sub-sample of 821 includes galaxies with a wide range of signal-to-noise ratios, star formation histories and even wavelength range (mainly due to each galaxy having different masks applied to it, according to the quality of the spectroscopic data). Figure 19 shows the number of recovered non-zero parameters in the sample, using VESPA. As an average, it falls below the synthetic examples studied in

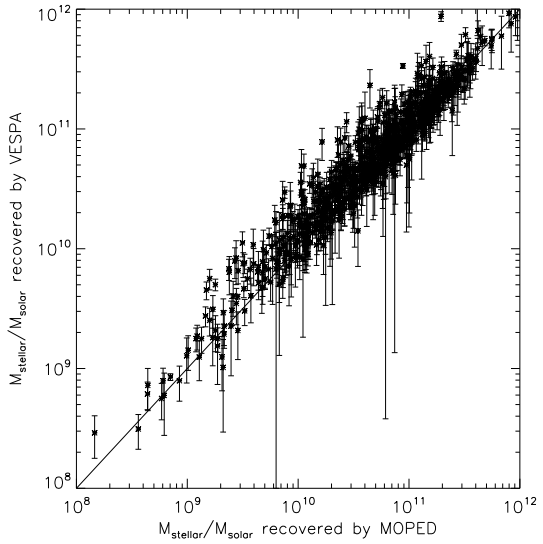

Figure 18. Galaxy stellar mass (in units of solar masses) as recovered by VESPA and MOPED for a sub-sample of 821 SDSS galaxies. The small percentage of galaxies with significantly larger VESPA masses have large extinction. The difference is accounted for by the fact that MOPED and VESPA use different dust models.

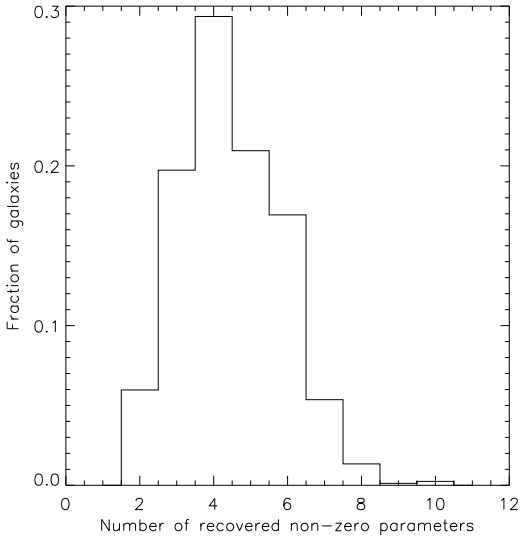

Figure 19. Number of non-zero parameters in solutions recovered from 821 SDSS galaxies with VESPA. Please note that these correspond to the *total* number of non-zero components in the solution vector $\mathbf{c}_\kappa$ and not to the number of recovered stellar populations. For a information about the number of recovered populations see Figure 14.

Section 3. This is not surprising, though, as each galaxy will have an unique and somewhat random combination of characteristics which will lead to a different number of parameters being recovered. The total combination of these sets of characteristics would be impossible to investigate using the empirical method described in Section 3, and here lies the advantage of VESPA of dynamically adapting to each individual case. Also important to note is the fact that the wavelength coverage is normally not continuous in an SDSS galaxy, due to masked regions. This was not modelled in Section 3, and is likely to further reduce the number of recovered parameters in any given case.

Perhaps more useful is to translate this number into a number of recovered significant stellar populations for each galaxy. We define a significant component as a stellar population which contributes 5 per cent or more to the total flux. Figure 20 shows the distribution of the number of significant components for our sub-sample of galaxies, as recovered by MOPED and VESPA. It is interesting to note that both methods recover on average a similar amount of components, even though MOPED has no explicit self-regularization mechanism, as VESPA clearly does.

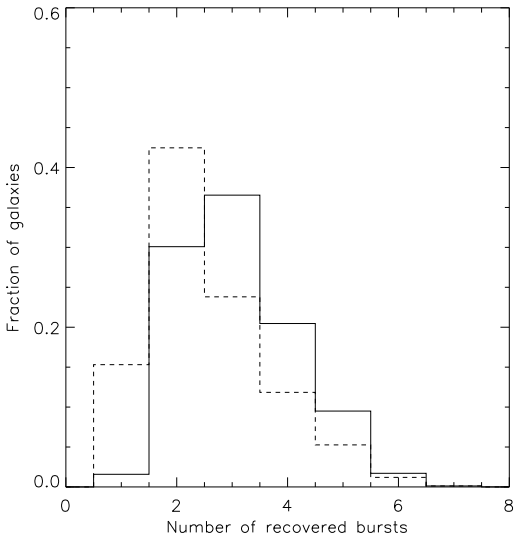

Figure 20. The distribution of the total number of recovered stellar populations which contribute 5 per-cent or more to the total flux of the galaxy, as recovered from MOPED (dashed line) and VESPA (solid line).

5 CONCLUSIONS

We have developed a new method to recover star formation and metallicity histories from integrated galactic spectra - VESPA. Motivated by the current limitations of other methods which aim to do the same, our goal was to develop an algorithm which is robust on a galaxy-by-galaxy basis. VESPA works with a dynamic parametrization of the

star formation history, and is able to adapt the number of parameters it attempts to recover from a given galaxy according to its spectrum. In this paper we tested VESPA against a series of idealised synthetic situations, and against SDSS data by comparing our results with those obtained with the well-established code, MOPED.

Using synthetic data we found the quality and resolution of the recovered solutions varied with factors such as type of star formation history, noise in the data and wavelength coverage. In the vast majority of cases, and within the estimated errors and bin-correlations, we observed a reliable reproduction of the input parameters. As the signal-to-noise decreases, it becomes increasingly difficult to recover robust solutions. Whereas our method cannot guarantee a perfect solution, we have shown that the self-regularization we imposed helped obtain a cleaner solution in an overwhelming majority of the cases studied.

On the real data analysis, we have studied possible effects from systematics using duplicate observations of the same set of galaxies, and have also compared VESPA's to MOPED's results obtained using the same data sample. We found that in the majority of cases our results are robust to possible systematics effects, but that in certain cases and particularly when calculating stellar masses, VESPA might underestimate the mass errors. However, we found no systematic bias in any of our tests. We have also shown that VESPA's results are in good agreement with those of MOPED for the same sample of galaxies. VESPA and MOPED are two fundamentally different approaches to the same problem, and we found good agreement both in a global sense by looking at the average star formation history of the sample, and in an individual basis by looking at the recovered stellar masses of each galaxy. VESPA typically recovered between 2 to 5 stellar populations from the SDSS sample.

VESPA's ability to adapt dynamically to each galaxy and to extract only as much information as the data warrant is a completely new way to tackle the problem of extracting information from galactic spectra. Our claim is that, for the most part, VESPA's results are robust for any given galaxy, but our claim comes with two words of caution. The first one concerns very noisy galaxies - in extreme cases (SNR$\approx$10 or less, at a resolution of 3$\mathring{A}$), it becomes very difficult to extract any meaningful information from the data. This uncertainty is evident in the large error bars and bin-correlations, and the solutions can be essentially unconstrained even at low-resolutions. We are therefore limited when it comes to analysing individual high-noise galaxies, which is the case of many SDSS objects. Our second word of caution concerns the stellar models used to analyse real galaxies - any method can only do as well as the models it bases itself upon. We are limited in our knowledge and ability to reproduce realistic synthetic models of stellar populations, and this is inevitably reflected in the solutions we obtain by using them. On the plus side, VESPA works with any set of synthetic models and can take advantage of improved versions as they are developed.

VESPA is fast enough to use on large spectroscopic samples (a typical SDSS galaxy takes 1 minute on an average workstation), and we are in the process of analysing SDSS's Data Release 5 (DR5), which consists of roughly half a million galaxies. Our first aim is to publish and exploit a catalogue of robust star formation histories, which we hope will be a valuable resource to help constrain models of galaxy formation and evolution.

6 ACKNOWLEDGMENTS

We are grateful to the anonymous referee for a very thoughtful report which led to material improvements in the paper.

RT is funded by the Fundação para a Ciência e a Tecnologia under the reference PRAXIS SFRH/BD/16973/04. RJ's research is supported by the NSF through grant PIRE-0507768 and AST-0408698 to the Atacama Cosmology Telescope.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/. The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

REFERENCES

Abazajian K., et al., 2005, AJ, 129, 1755
Abraham R. G., et al., 2007, preprint (astro-ph/0701779)
Alongi M., Bertelli G., Bressan A., Chiosi C., Fagotto F., Greggio L., Nasi E., 1993, A&AS, 97, 851
Barber T., Meiksin A., Murphy T., 2006, preprint (astro-ph/0611053)
Bressan A., Fagotto F., Bertelli G., Chiosi C., 1993, A&AS, 100, 647
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Bundy K., et al., 2006, ApJ, 651, 120

Chabrier G., 2003, PASP, 115, 763
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Cid Fernandes R., Asari N. V., Sodré L., Stasińska G., Mateus A., Torres-Papaqui J. P., Schoenell W., 2007, MNRAS, 375, L16
Cid Fernandes R., Gu Q., Melnick J., Terlevich E., Terlevich R., Kunth D., Rodrigues Lacerda R., Joguet B., 2004, MNRAS, 355, 273
Erb D. K., Steidel C. C., Shapley A. E., Pettini M., Reddy N. A., Adelberger K. L., 2006, ApJ, 647, 128
Fagotto F., Bressan A., Bertelli G., Chiosi C., 1994a, A&AS, 104, 365
Fagotto F., Bressan A., Bertelli G., Chiosi C., 1994b, A&AS, 105, 29
Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, MNRAS, 362, 41
Girardi L., Bressan A., Chiosi C., Bertelli G., Nasi E., 1996, A&AS, 117, 113
Glazebrook K., et al., 2003, ApJ, 587, 55
Gordon K. D., Clayton G. C., Misselt K. A., Landolt A. U., Wolff M. J., 2003, ApJ, 594, 279
Heavens A., Panter B., Jimenez R., Dunlop J., 2004, Nat, 428, 625
Heavens A. F., Jimenez R., Lahav O., 2000, MNRAS, 317, 965
Hopkins A. M., Connolly A. J., Szalay A. S., 2000, AJ, 120, 2843
Kauffmann G., et al., 2003, MNRAS, 341, 33
Kennicutt Jr. R. C., 1998, ApJ, 498, 541
Lawson C., Hanson R., 1974, Solving Least Squares Problems. Prentice-Hall, Inc.
Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, MNRAS, 283, 1388
Mathis H., Charlot S., Brinchmann J., 2006, MNRAS, 365, 385
Noeske K. G., et al., 2007, preprint (astro-ph/0703056)
Ocvirk P., Pichon C., Lançon A., Thiébaut E., 2006, MNRAS, 365, 46
Panter B., Heavens A. F., Jimenez R., 2003, MNRAS, 343, 1145
Panter B., Jimenez R., Heavens A. F., Charlot S., 2006, preprint (astro-ph/0608531)
Spergel D. N., et al., 2003, ApJ Supplement Series, 148, 175
Stark P. B., Parker R. L., 1995, Computational Statistics, 10, 143
Strauss M. A., et al., 2002, AJ, 124, 1810
Thomas D., Maraston C., Bender R., 2003, MNRAS, 339, 897
Tremonti C. A., et al., 2004, ApJ, 613, 898
Verma A., Lehnert M. D., Foerster Schreiber N. M., Bremer M. N., Douglas L., 2007, preprint (astro-ph/0701725)
Worthey G., 1994, ApJ Supplement Series, 95, 107
York D. G., et al., 2000, AJ, 120, 1579